



xiaomi

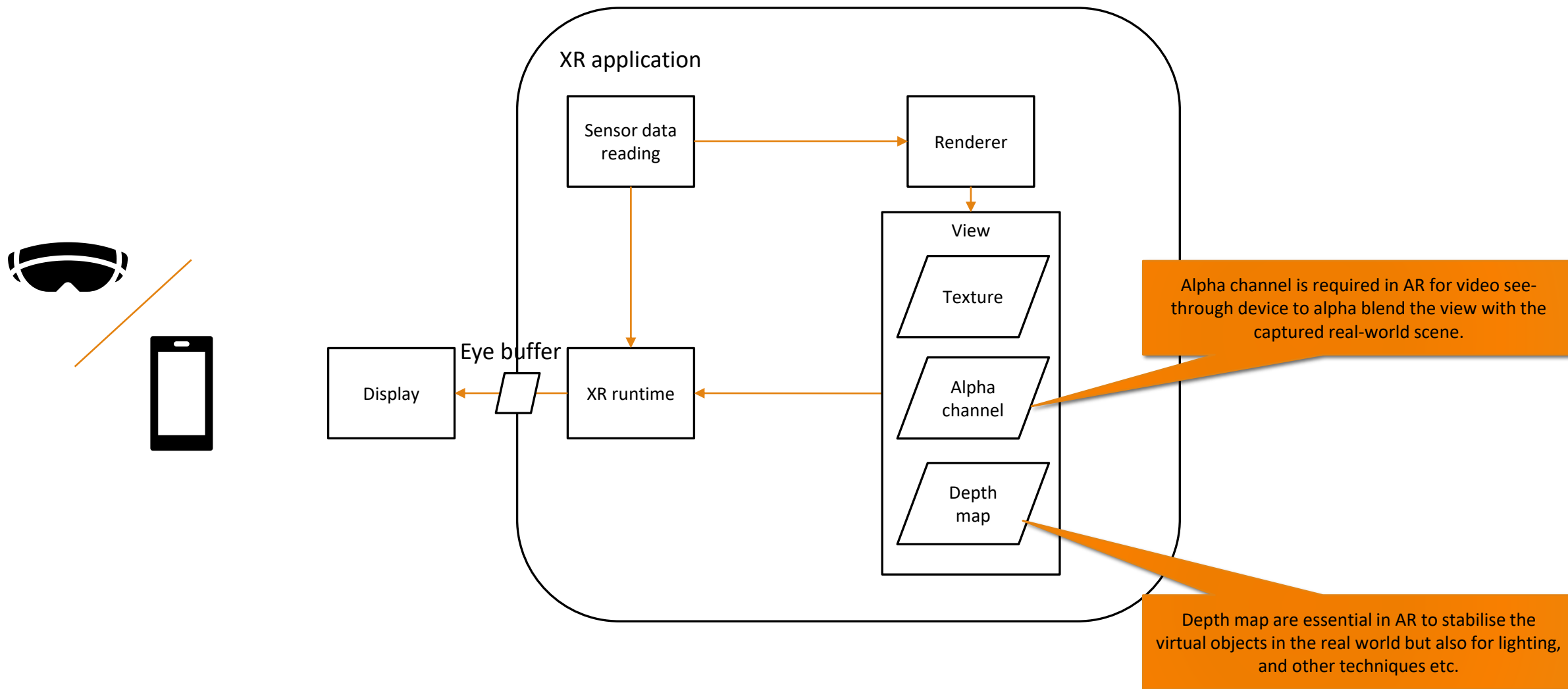
Carriage of depth and alpha maps as HEVC single-layer bitstreams

JVET-AG0144-v2

E. THOMAS, E. POTETSIANAKIS, E. ALEXIOU, R.
GHAZNAVI-YOUVALARI, M. ABDOLI, M-L. CHAMPEL
01/17/2024

- 1. Background: HEVC-based split rendering service (3GPP SA4)**
- 2. Design goals for depth and alpha maps in HEVC single-layer bitstreams**
- 3. Possible solutions for each design goal**

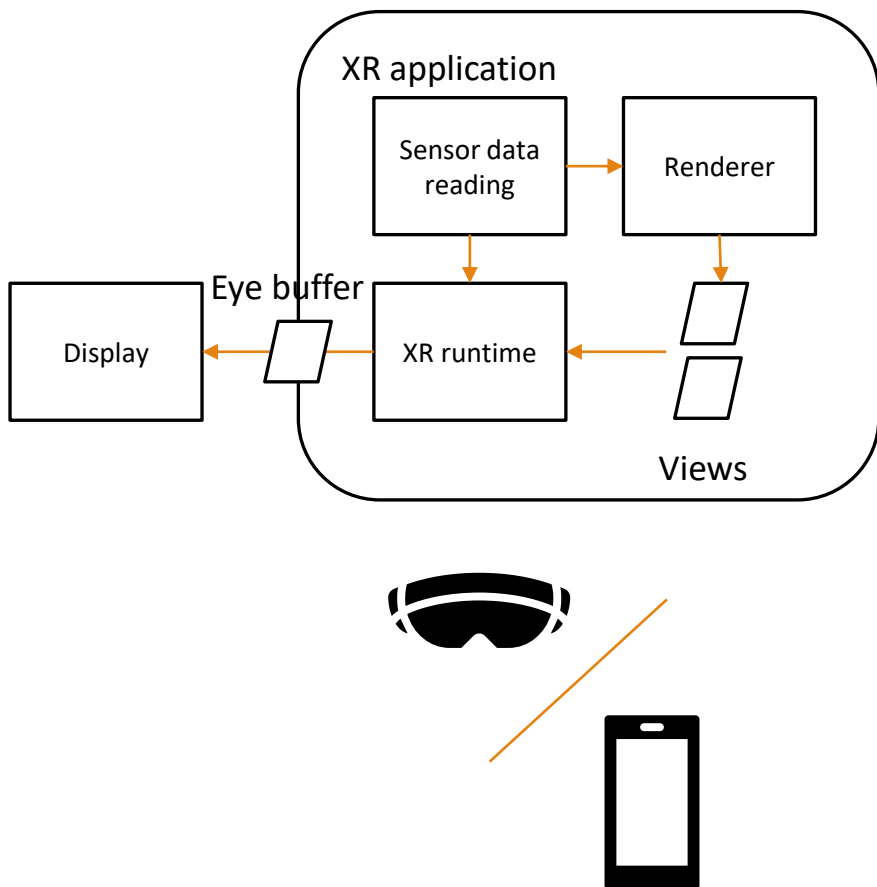
1. Background: XR rendering



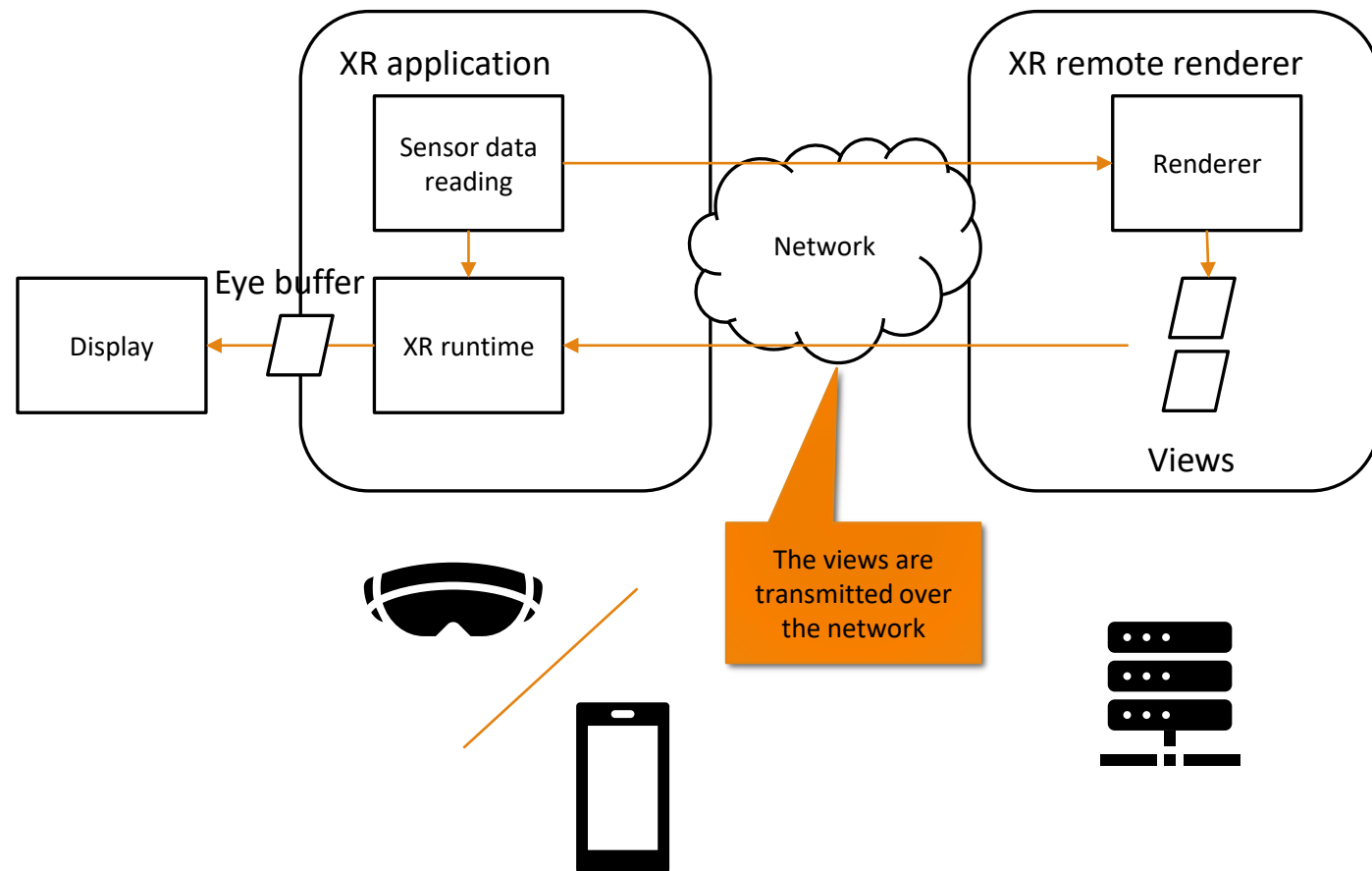
More information can be found in the OpenXR specification ([link](#))

1. Background: Split rendering concept

Local XR rendering



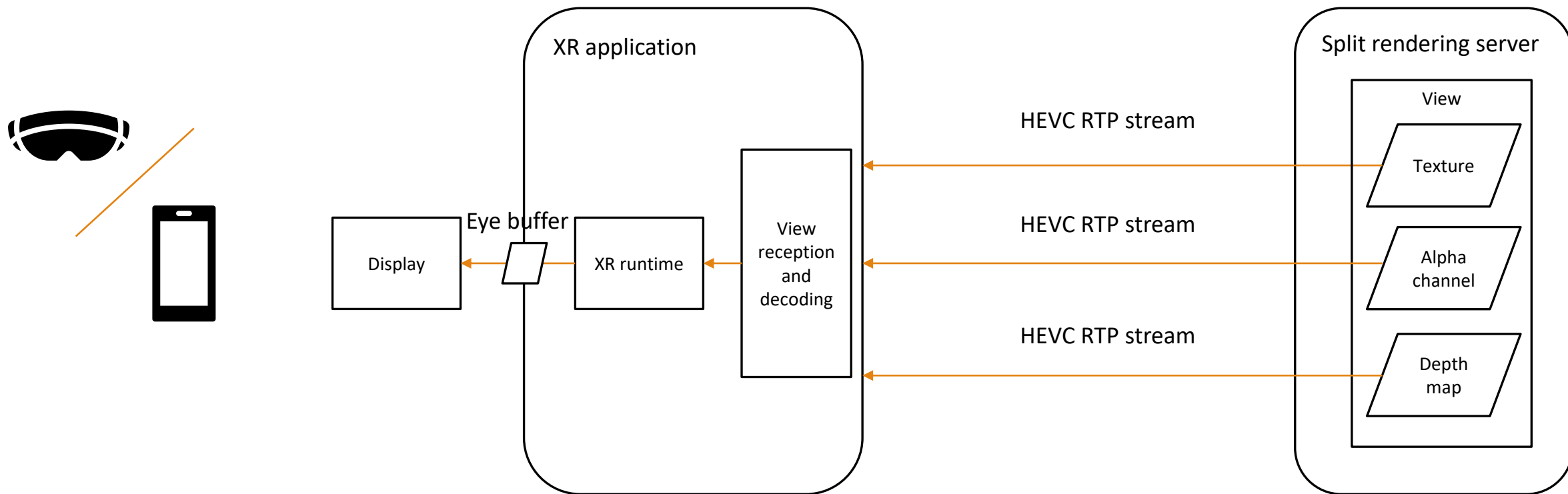
XR split rendering



1. Background: 3GPP HEVC split rendering



Overview of 3GPP SA4 split rendering service*



* See 3GPP SA4 TS 26.565 Split Rendering Media Service Enabler ([link](#)) for more details

2. Design goals

- **Enabling depth and alpha maps in HEVC single-layer bitstream by signalling:**
 - 1. Indication that the bitstream carries a depth map sequence or an alpha map sequence.**
 - 2. Indication on the mapping used between the depth/alpha map (mono-channel) and the luma and chroma components.**
 - 3. Information specifying how to interpret the sample values from the decoded pictures to reconstruct the depth/alpha maps.**

1. Information indicating that the bitstream carries a depth map sequence or an alpha map sequence.

Option 1.1

Defining a content type parameter to identify depth map and alpha map sequences in a new SEI message.

content_type_information(payloadSize) {	Descriptor
cti_content_tpe_idc	u(4)
cti_component_mapping_idc	u(8)
}	

cti_content_type_idc	Content type description
0	The current layer represents a depth map sequence.
1	The current layer represents an alpha map sequence.
2-10	Reserved
11-15	Unspecified and left for usage by external specifications.

Option 1.2

Adding colour_primaries code points in the VUI.

Value	Primaries	Informative remark
0	Reserved	For future use by ITU-T ISO/IEC
1	primary x y green 0.300 0.600 blue 0.150 0.060 red 0.640 0.330 white D65 0.312 7 0.329 0	Rec. ITU-R BT.709-6 Rec. ITU-R BT.1361-0 conventional colour gamut system and extended colour gamut system (historical) IEC 61966-2-1 sRGB or sYCC IEC 61966-2-4 SMPTE RP 177 (1993) Annex B
...		
23	Depth map	This signal is a depth picture
24	Alpha map	This signal is an alpha map
25..255	Reserved	For future use by ITU-T ISO/IEC

Note: See Word contribution in JVET-AG0144 for entire spec text

3. Possible solutions for design goal #2

2. Indication on the mapping used between the depth/alpha map (mono-channel) and the luma and chroma components.

Option 2.1

Defining a component mapping parameter in the new SEI message.

content_type_information(payloadSize) {	Descriptor
<code>cti_content_tpe_idc</code>	<code>u(4)</code>
<code>cti_component_mapping_idc</code>	<code>u(8)</code>
}	

Value	Description
0	This value defines a one-to-one mapping between the luma component and the content image.
1	This value defines a one-to-one mapping between the luma component and the content image. If present, the chroma components are to be discarded and their sample values are set to 1 << (BitDepthC – 1).
2	<p>This value defines a mapping in which the three components (luma and two chroma) constitute different bit ranges of the content image sample values. The coded picture format is 4:4:4 which provides a bit precision of BitDepth + 2*BitDepthC bits, i.e. the bit depth of the luma component plus two times the bit depth of the chroma component. The mapping is defined as follows:</p> <p>Y: $y(i,j) = (m(i,j) \& 0xFF0000) \gg 16$</p> <p>Cb: $cb(i,j) = (m(i,j) \& 0x00FF00) \gg 8$</p> <p>Cr: $cr(i,j) = m(i,j) \& 0x0000FF$</p> <p>With $y(i,j)$, $cb(i,j)$ and $cr(i,j)$, respectively the samples of the Y, Cb and Cr components and $m(i,j)$ the sample of the content image.</p>
3..255	Reserved

Note: See Word contribution in JVET-AG0144 for entire spec text

3. Possible solutions for design goal #2

2. Indication on the mapping used between the depth/alpha map (mono-channel) and the luma and chroma components.

Option 2.2

Adding matrix_coeffs code points for mono-channel (luma only) in the VUI.

Value	Matrix	Informative remark
...
18	1st colour component	A monochromatic representation with only the first colour component in the original representation present in the coded image. Typically used for Green only; however, may also be used for the Y tristimulus value (luminance), when the colour primaries are set to XYZ, or to indicate that only the first component of the decoded image shall be interpreted for images not representing texture.
19	2nd colour component	A monochromatic representation with only the second colour component in the original representation present in the coded image. Typically used for Red only; however, may also be used for the X tristimulus value, when the colour primaries are set to XYZ, or to indicate that only the first component of the decoded image shall be interpreted for images not representing texture.
20	3rd colour component	A monochromatic representation with only the third colour component in the original representation present in the coded image. Typically used for Blue only; however, may also be used for the Z tristimulus value, when the colour primaries are set to XYZ, or to indicate that only the first component of the decoded image shall be interpreted for images not representing texture.
21	Luma only KR = 0.2126; KB = 0.0722	A monochromatic representation of the luma component of the Rec. ITU-R BT.709-6 representation.
22	Luma only KR = 0.2627; KB = 0.0593	A monochromatic representation of the luma component of the Rec. ITU-R BT.2020-2 non-constant luminance representation.
23	I only - ICTCP	A monochromatic representation of the I component of ICTCP representation.
24..255	Reserved	For future use by ITU-T ISO/IEC

Note: See Word contribution in JVET-AG0144 for entire spec text

3. Possible solutions for design goal #3

3. Information specifying how to interpret the sample values from the decoded pictures to reconstruct the depth/alpha maps.

Option 3.1

New SEI messages based on the multi-layer extension but defined for the single layer context (no reuse of annexes F and G).

depth_map_info(payloadSize) {	Descriptor
z_near_flag	u(1)
z_far_flag	u(1)
depth_representation_type	ue(v)
if(z_near_flag)	
depth_rep_info_element(ZNearSign, ZNearExp, ZNearMantissa, ZNearManLen)	
if(z_far_flag)	
depth_rep_info_element(ZFarSign, ZFarExp, ZFarMantissa, ZFarManLen)	
}	

■ The differences with the existing depth and alpha SEI:

- No disparity mode for depth since two views are needed
- The alpha channel SEI syntax could be reused as is
- Allowed in a non-auxiliary layer (e.g. single layer)

depth_representation_type	Interpretation
0	Each decoded luma sample value of a picture represents an inverse of Z value that is uniformly quantized into the range of 0 to maxVal, inclusive. When z_far_flag is equal to 1, the luma sample value equal to 0 represents the inverse of ZFar (specified below). When z_near_flag is equal to 1, the luma sample value equal to maxVal represents the inverse of ZNear (specified below).
1	Each decoded luma sample value of a picture represents a Z value uniformly quantized into the range of 0 to maxVal, inclusive. When z_far_flag is equal to 1, the luma sample value equal to 0 corresponds to ZFar (specified below). When z_near_flag is equal to 1, the luma sample value equal to maxVal represents ZNear (specified below).
2	Each decoded luma sample value of a picture represents an Euclidean distance from the camera to the object in the scene.
3.15	Reserved

Note: See Word contribution in JVET-AG0144 for entire spec text

3. Possible solutions for design goal #3

3. Information specifying how to interpret the sample values from the decoded pictures to reconstruct the depth/alpha maps.

Option 3.2

Modifications of existing SEI messages to be used in the single layer context.

■ Example changes in the depth representation SEI message:

The syntax elements in the depth representation information SEI message specify various parameters for auxiliary pictures of type AUX_DEPTH **or non-auxiliary pictures carrying depth maps**. Those parameters are signalled for the purpose [...]

When present, the depth representation information SEI message shall be associated with **either:**

- one or more layers with AuxId value equal to AUX_DEPTH
- **a layer with primary pictures carrying depth maps**

d_min_flag equal to 0 specifies that the syntax elements specifying the minimum disparity value are not present in the syntax structure. d_min_flag equal to 1 specifies that the syntax elements specifying the minimum disparity value are present in the syntax structure.

When associated with primary pictures carrying depth maps, it is a requirement of bitstream conformance that d_min_flag shall be equal to 0.
[...]

Note: See Word contribution in JVET-AG0144 for entire spec text

- Split rendering is critical for XR applications.
 - 3GPP SA4 is developing a split rendering service specification.
 - This service requires multiple RTP HEVC streams for texture sequence and depth and alpha map sequences.
 - Alternative solutions are proposed to enable HEVC single layer carriage of depth and alpha maps.
-
- We recommend documenting the preferred approach from the proposed alternatives in the relevant TuC documents.